

The time of graph bootstrap percolation

Karen Gunderson[†], Sebastian Koch, and Michał Przykucki^{*}

K. Gunderson
Department of Mathematics
University of Manitoba
Winnipeg, MB R3T 2N2
Canada
e-mail: karen.gunderson@umanitoba.ca

S. Koch
Department of Pure Mathematics and Mathematical Statistics
University of Cambridge
Cambridge CB3 0WB
United Kingdom
e-mail: sk629@cam.ac.uk

M. Przykucki
St Anne's College
University of Oxford
Oxford OX2 6HS
United Kingdom
e-mail: przykucki@maths.ox.ac.uk

Abstract: Graph bootstrap percolation, introduced by Bollobás in 1968, is a cellular automaton defined as follows. Given a “small” graph H and a “large” graph $G = G_0 \subseteq K_n$, in consecutive steps we obtain G_{t+1} from G_t by adding to it all new edges e such that $G_t \cup e$ contains a new copy of H . We say that G percolates if for some $t \geq 0$, we have $G_t = K_n$.

For $H = K_r$, the question about the size of the smallest percolating graphs was independently answered by Alon, Frankl and Kalai in the 1980's. Recently, Balogh, Bollobás and Morris considered graph bootstrap percolation for $G = G(n, p)$ and studied the critical probability $p_c(n, K_r)$, for the event that the graph percolates with high probability. In this paper, using the same setup, we determine, up to a logarithmic factor, the critical probability for percolation by time t for all $1 \leq t \leq C \log \log n$.

AMS 2000 subject classifications: primary 60K35; secondary 60C05.

Keywords and phrases: bootstrap percolation, weak saturation.

1. Introduction

Cellular automata, introduced by von Neumann [19] after a suggestion of Ulam [21], are dynamical systems acting on graphs using local and homogeneous update rules. The H -bootstrap percolation process is one example of such an automaton and can be described as follows. Given a fixed graph H and a graph $G \subseteq K_n$, set $G_0 = G$ and then, for each $t = 0, 1, 2, \dots$, let

$$G_{t+1} = G_t \cup \{e \in E(K_n) : \exists H \text{ with } e \in H \subseteq G_t \cup e\}. \quad (1)$$

Let $\langle G \rangle_H = \bigcup_{t=0}^{\infty} G_t$ denote the *closure* of G under H -bootstrap percolation. We say that G *percolates* in the H -bootstrap process (or H -percolates), if $\langle G \rangle_H = K_n$. (See Figure 1).

The notion of H -percolation, introduced by Bollobás in 1968 [7] under the name of *weak saturation*, has been extensively studied in the case where H is a complete graph. Initially, the extremal properties of the H -bootstrap process attracted the most attention. Alon [1], Frankl [15] and Kalai [18] independently confirmed a conjecture of Bollobás and proved that the smallest K_r -percolating

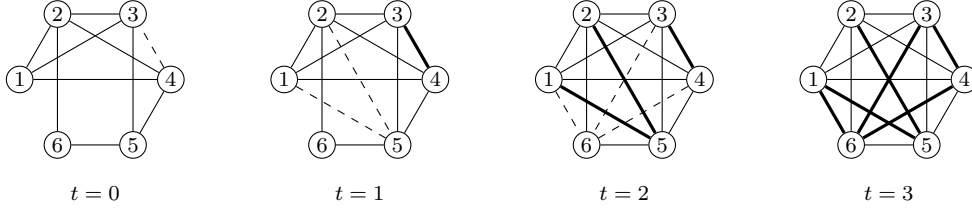


FIG 1. An example of the K_4 -bootstrap percolation process. Dashed edges are added to the graph on the next time step.

graphs on n vertices have size $\binom{n}{2} - \binom{n-r+2}{2}$.

Recently, Bollobás [9] observed a strong connection between weak saturation and r -neighbour bootstrap percolation, a dynamical process suggested in 1979 by Chalupa, Leath and Reich [13]. For an integer $r \geq 2$, the r -neighbour bootstrap process on a graph $G = (V, E)$ with an ‘initial set’ of vertices $A \subseteq V$ is defined by setting $A_0 = A$ and for $t = 0, 1, 2, \dots$, defining

$$A_{t+1} = A_t \cup \{v \in V : |N(v) \cap A_t| \geq r\}, \quad (2)$$

where $N(v)$ is the set of neighbours of v in G . The set $\langle A \rangle = \bigcup_{t=0}^{\infty} A_t$ is the closure of A and we say that A percolates if $\langle A \rangle = V$. Often, the vertices in the set A_t are called ‘infected’ and the remaining vertices are ‘healthy’. The usual question asked in the context of r -neighbour bootstrap percolation is the following: if the vertices of G are initially infected independently at random with probability p , for what values of p is percolation likely to occur? The probability of percolation is clearly non-decreasing in p hence it is natural to define the *critical probability* $p_c(G, r)$ as

$$p_c(G, r) = \inf\{p : \mathbb{P}_p(\langle A \rangle = V(G)) \geq 1/2\}. \quad (3)$$

The study of critical probabilities has brought numerous and often very sharp results for various graphs G and the values of the infection threshold. For example, van Enter [14] and Schonmann [20] studied r -neighbour bootstrap percolation on \mathbb{Z}^d , Holroyd [16], Balogh, Bollobás and Morris [4], Balogh, Bollobás, Duminil-Copin and Morris [3] analysed finite grids, while Balogh and Pittel [6], Janson, Łuczak, Turova and Vallier [17] and Bollobás, Gunderson, Holmgren, Janson and Przykucki [10] worked with random graphs.

Motivated by this approach, Balogh, Bollobás and Morris [5] defined the critical probability for H -bootstrap percolation on K_n to be

$$p_c(n, H) = \inf\{p : \mathbb{P}_p(\langle G_{n,p} \rangle_H = K_n) \geq 1/2\}, \quad (4)$$

where $G_{n,p}$ is the Erdős-Rényi random graph, obtained by choosing every edge of K_n independently at random with probability p . In [5], they showed that for all $r \geq 4$, taking $\lambda(r) = (\binom{r}{2} - 2)/(r - 2)$ and $n \in \mathbb{N}$ sufficiently large,

$$\frac{n^{-1/\lambda(r)}}{2e \log n} \leq p_c(n, K_r) \leq n^{-1/\lambda(r)} \log n. \quad (5)$$

In this paper we focus on a different question related to K_r -bootstrap percolation. Namely, for what values of p is percolation likely to occur *by time* t ? Defining K_r -bootstrap percolation as in (1), let $T = T(n, r, G_0) = \min\{t : G_t = K_n \text{ in the } K_r\text{-bootstrap process}\}$. Define the *critical probability for percolation by time* t to be

$$p_c(n, r, t) = \inf\{p : \mathbb{P}_p(T \leq t) \geq 1/2\}. \quad (6)$$

*Supported in part by EU project MULTIPLEX no. 317532.

†Research carried out while affiliated with the Heilbronn Institute for Mathematical Research at the University of Bristol.

For notational convenience, for any $r \geq 4$ and any $t \geq 1$, set

$$\tau = \tau(r) = \binom{r}{2} - 1, \quad e_t = \tau^t, \quad \text{and} \quad v_t = (r-2) \frac{\tau^t - 1}{\tau - 1} + 2. \quad (7)$$

The following theorem is the main result of this paper.

Theorem 1.1. Let $r \geq 4$ and $t = t(n) \leq \frac{\log \log n}{3 \log \tau}$. Let $(p_n)_{n=1}^\infty$ be a sequence of probabilities, let $\omega(n) \rightarrow \infty$ and let $T = T(n, r, G_{n, p_n})$. Under the K_r -bootstrap process,

- (i) if, for all n , $p_n \geq n^{-(v_t-2)/e_t} \log n$, then $\mathbb{P}_{p_n}(T \leq t) \rightarrow 1$ as $n \rightarrow \infty$ and
- (ii) if, for all n , $p_n \leq n^{-(v_t-2)/e_t} / \omega(n)$, then $\mathbb{P}_{p_n}(T \leq t) \rightarrow 0$ as $n \rightarrow \infty$.

Thus, Theorem 1.1 shows that for all $r \geq 4$ and $1 \leq t \leq \frac{\log \log n}{3 \log \tau}$, and $\omega(n) \rightarrow \infty$, for n sufficiently large we have

$$\frac{n^{-(v_t-2)/e_t}}{\omega(n)} \leq p_c(n, r, t) \leq n^{-(v_t-2)/e_t} \log n. \quad (8)$$

Similar questions related to the time of r -neighbour bootstrap percolation on grids have recently been studied by Bollobás, Holmgren, Smith and Uzzell [11], Bollobás, Smith and Uzzell [12] and by Balister, Bollobás and Smith [2]. The time of the r -neighbour bootstrap process on the random graph $G(n, p)$ was analysed in [17].

Before we continue, let us briefly discuss how the bounds on the critical probability for percolation by time t in Theorem 1.1 relate to the bounds on $p_c(n, K_r)$ in (5). By the definitions of $p_c(n, K_r)$ in (4) and of $p_c(n, r, t)$ in (6), we clearly have $p_c(n, K_r) \leq p_c(n, r, t)$. Ignoring the polylogarithmic factors in the critical probability, we see that the exponent of n in (8) can be re-written using the identities in (7) as follows, writing $\lambda = \lambda(r)$,

$$-\frac{v_t - 2}{e_t} = -\frac{(r-2) \frac{\tau^t - 1}{\tau - 1}}{\tau^t} = -\frac{r-2}{\binom{r}{2} - 2} \frac{\tau^t - 1}{\tau^t} = -\frac{1}{\lambda} + \frac{1}{\lambda \tau^t}. \quad (9)$$

For $t = \frac{\log \log n}{3 \log \tau}$, which is the maximum value of t covered by Theorem 1.1, in (9) we obtain

$$-\frac{1}{\lambda} + \frac{1}{\lambda \tau^t} = -\frac{1}{\lambda} + \frac{1}{\lambda \tau^{(\log \log n)/3}} = -\frac{1}{\lambda} + \frac{1}{\lambda (\log n)^{1/3}}.$$

Hence the exponent of n in (8) tends to the exponent in (5) for $t = \frac{\log \log n}{3 \log \tau}$. In fact, if one managed to show that (8) holds all the way up to $t = \frac{\log \log n}{\log \tau}$ then the bounds in (5) and (8) would match up to polylogarithmic factors.

The proofs of both statements of Theorem 1.1 rely on the properties of a family of graphs, denoted $\{F_t : t \geq 1\}$, that are described in detail in Section 3. For each t , there is a pair of vertices in $V(F_t)$ so that if F_t occurs as a subgraph of G_0 , then that pair is guaranteed to be added to the graph by time t . The graph F_t is thought of as ‘anchored’ on that special pair of vertices.

To prove Statement (i) of Theorem 1.1, Janson’s inequality is used to bound from below the probability that a particular pair $\{x, y\}$ is contained as the anchor vertices in some copy of F_t . To establish a bound in this way, estimates are needed on the probability that two overlapping copies of F_t occur in G_0 . This amounts to determining the minimum possible ratio of edges to vertices in some non-trivially overlapping pair. It turns out that the minimum ratio is not obtained for one of the extreme cases, i.e., neither for two copies of F_t that share only one vertex, nor for two copies that share all but one vertex. Even though we do not prove it directly in this paper, the minimal density of two overlapping copies of F_t is not monotone in the size of their common part and it can be shown that, as $t \rightarrow \infty$, the two overlapping copies of F_t that minimise the edge-to-vertex ratio share an approximately $4/((r+1)(r-2))$ proportion of the vertex set. Bounding this ratio from below for all

possible configurations of two such copies is the main challenge in the proof of the upper bound on $p_c(n, r, t)$, and is dealt with in detail in Section 3.1.

To prove Statement (ii) of Theorem 1.1 we employ two extremal results about graphs that add an edge e to the graph in at most t time steps: one of them to bound the number of their vertices from above, and one (a corollary of a highly nontrivial result in [5]) to bound their edge density from below. Then, for p as in Statement (ii) of Theorem 1.1, we show that with high probability no such graph can be found in $G_{n,p}$. This completes the proof of our main result.

The remaining sections of the paper are organised as follows. In Section 2 we briefly discuss the K_3 -bootstrap percolation process which behaves differently than K_r -bootstrap processes when $r \geq 4$. In Section 3, we introduce the graphs F_t that are the main focus of the proofs to come and prove some key properties. In Section 3.1, which is the crucial part of our argument, we prove some properties of graphs consisting of two overlapping copies of F_t . In Sections 4 and 5 we prove Statements (i) and (ii) of Theorem 1.1 respectively. Finally, in Section 6, some open problems are stated.

2. K_3 -bootstrap percolation

In this section we discuss the special case of the K_3 -bootstrap process. Observe that a graph G percolates in K_3 -bootstrap percolation if and only if G is connected. Also, at every time step each non-edge between vertices at distance 2 is added to the graph. Therefore, if G is a connected graph with diameter d , then the diameter of the graph obtained from G after one step of the K_3 -bootstrap process is $\lceil d/2 \rceil$. Hence, G percolates in $\lceil \log_2 d \rceil$ time steps.

The diameter of random graphs was investigated by Bollobás [8] who proved the following theorem.

Theorem 2.1. Let $G_{n,p}$ be the Erdős-Rényi random graph.

1. Suppose $p^2 n - 2 \log n \rightarrow \infty$ and $n^2(1-p) \rightarrow \infty$. Then $G_{n,p}$ has diameter 2 whp.
2. Suppose the functions $d = d(n) \geq 3$ and $0 < p = p(n) < 1$ satisfy $(\log n)/d - 3 \log \log n \rightarrow \infty$, $p^d n^{d-1} - 2 \log n \rightarrow \infty$ and $p^{d-1} n^{d-2} - 2 \log n \rightarrow -\infty$. Then $G_{n,p}$ has diameter d whp.

In order to re-phrase this result in the form of intervals for p in which the diameter is constant with high probability, let $\omega(n) = o(\log n)$ tend to infinity arbitrarily slowly. Clearly, if $p \geq 1 - 1/(n^2 \omega(n))$ then whp. $G_{n,p} = K_n$ which has diameter 1. Simplifying a bit, Theorem 2.1 implies that if

$$\sqrt{\frac{2 \log n + \omega(n)}{n}} \leq p \leq 1 - \frac{\omega(n)}{n^2}$$

then $G_{n,p}$ has diameter 2, and that for $3 \leq d \leq \log n / 4 \log \log n$, if

$$p(n) \in \left((2 \log n + \omega(n))^{\frac{1}{d}} n^{-\frac{d-1}{d}}, (2 \log n - \omega(n))^{\frac{1}{d-1}} n^{-\frac{d-2}{d-1}} \right)$$

then the random graph $G_{n,p(n)}$ has diameter d with high probability. This answers the question about the time of K_3 -bootstrap percolation.

3. Adding an edge to the graph using sparse subgraphs

Throughout the following sections, fix $r \geq 4$. As r is fixed, for simplicity, it is often omitted from the notation. We define a family $\{F_t : t \geq 1\}$ of graphs that add a given pair as an edge to the graph in the K_r -bootstrap process exactly at time t . We prove that these are the “sparsest” minimal such graphs (i.e., they minimise the ratio of the number of edges to the number of vertices). The main result in this section, presented in Section 3.1, is a lower bound on the edge-density of two non-disjoint copies of the graph F_t . This bound is the key element of arguments for the proof of Statement (i) of Theorem 1.1.

The graph F_t is defined recursively and the fixed edge that is added to the graph at time t using F_t will always be denoted by $e_0 = \{1, 2\}$.

For $t = 1$, set $F_1 = K_r - e_0$, an r -clique missing one edge.

For each $t \geq 1$, given F_t , for each $e \in E(F_t)$, let $V(e)$ be a set of $r - 2$ new vertices and let $K(e)$ be a copy of $K_r - e$, an r -clique missing one edge, on vertex set $V(e) \cup e$. Then, F_{t+1} is defined to be the graph with vertex set

$$V(F_{t+1}) = V(F_t) \cup \left(\bigcup_{e \in E(F_t)} V(e) \right)$$

and edge set

$$E(F_{t+1}) = \bigcup_{e \in E(F_t)} E(K(e))$$

(see Figure 2). Note that any edge $e \in E(F_{t+1})$ is incident to at least one vertex in $V(f)$ for some $f \in E(F_t)$, i.e., to a vertex in $V(F_{t+1}) \setminus V(F_t)$.

Recall that we define $\tau = \binom{r}{2} - 1$ and the numbers e_t and v_t in equation (7). By induction on t , it can be shown that for every $t \geq 1$, the number of edges and vertices in the graph F_t are given by

$$e_t = e(F_t) = |E(F_t)| = \tau^t \text{ and} \quad (10)$$

$$v_t = v(F_t) = |V(F_t)| = |V(F_{t-1})| + e_{t-1}(r - 2) = 2 + (r - 2) \frac{\tau^t - 1}{\tau - 1}. \quad (11)$$

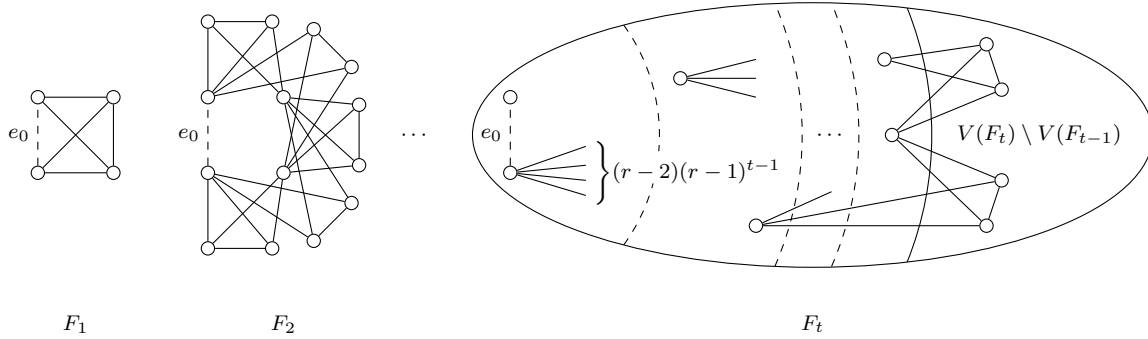


FIG 2. Construction of the graph F_t . Note that every edge in F_t is incident to at least one vertex in $V(F_t) \setminus V(F_{t-1})$.

Lemma 3.1. In the K_r -bootstrap process started from F_t the edge e_0 is added to the graph in exactly t steps.

Proof. We prove this lemma by induction on t . The statement is trivial for $t = 1$ as $F_1 = K_r - e_0$. Assume that the Lemma holds for $t = k \geq 1$. Note that after one step of the process started from F_{k+1} we obtain a copy of F_k in our graph since F_{k+1} is obtained from F_k by placing a copy of K_r minus an edge on every edge of F_k . Thus e_0 is added to the graph after at most $k + 1$ steps of the process started from F_{k+1} .

The construction of F_{k+1} can be also seen as placing a copy of F_k on each of the τ edges of $F_1 = K_r - e_0$. By induction we know that these copies of F_k on their own add the respective edges of F_1 , their *anchor* edges, in k time steps. This process could possibly accelerate if some interaction between two different copies of F_k occurred early in the process, say, before time k when the F_k 's on their own add their respective anchor edges. Therefore, let F^1, F^2, \dots, F^τ be these different copies of F_k in F_{k+1} . By construction of F_{k+1} we have that for all $i \neq j$:

1. F^i and F^j share at most one vertex,
2. no vertex of F^i other than the vertices in its anchor edge can have a neighbour outside F^i .

For every $s \in \{0, 1, \dots, k\}$, and $i \in \{1, 2, \dots, \tau\}$, let $E_{i,s}$ be the edges within $V(F^i)$ in the s -th step of the K_r -bootstrap process on the subgraph induced by F^i and let E_s be the edges in F_{k+1} in the s -th step of the K_r bootstrap process. The aim is to show that for $s \leq k$, there is no interaction between the processes in each F^i so that $E_s = \bigcup_{i=1}^{\tau} E_{i,s}$. This is proved by induction on s . Note that for $s \leq k-1$, if it is true that $E_s = \bigcup_{i=1}^{\tau} E_{i,s}$, then any triangle in the graph at that time is contained entirely within the vertex set of one of the F^i . Also, recall that $E_{i,s}$ contains only edges within $V(F^i)$.

For $s = 0$, the statement $E_0 = \bigcup_{i=1}^{\tau} E_{i,0}$ is true by construction.

Suppose that for some $s \in [0, k-1]$, that $E_s = \bigcup_{i=1}^{\tau} E_{i,s}$ and that there is an edge $\{w_1, w_2\} \in E_{s+1} \setminus \bigcup_{i=1}^{\tau} E_{i,s+1}$. As the edge $\{w_1, w_2\}$ must have been added at time $s+1$, let w_3, \dots, w_r be the $r-2$ other vertices in the copy of K_r minus an edge that witness $\{w_1, w_2\}$ being added at time $s+1$. Then, at time s , for any $i, j > 2$, the vertices w_1, w_i, w_j form a triangle, as do the vertices w_2, w_i, w_j . Then, by the induction hypothesis, there is an i_0 so that $w_1, w_2, w_i, w_j \in V(F^{i_0})$. Again, by hypothesis, this means that the edges $\{w_1, w_i\}, \{w_2, w_i\}, \{w_1, w_j\}, \{w_2, w_j\}, \{w_i, w_j\}$ are all contained in $E_{i_0,s}$. This is a contradiction, as then the edge $\{w_1, w_2\}$ would have been added within the graph F^{i_0} at time $s+1$. This completes the induction on s .

Thus the first k steps of the process started from F_{k+1} look like τ independent processes started from F_k 's. Hence e_0 is not added to the graph by time k . This completes the proof of the lemma. \square

Recall that we denote

$$\lambda = \frac{\binom{r}{2} - 2}{r-2} = \frac{r+1}{2} - \frac{1}{r-2} \quad (12)$$

and that τ can be written as

$$\tau = \binom{r}{2} - 1 = \frac{(r+1)(r-2)}{2}. \quad (13)$$

Note that (12) and (13) imply that

$$(r-2)\lambda = \tau - 1. \quad (14)$$

Define for convenience,

$$c_t = \frac{1}{\tau^t - 1}, \quad (15)$$

which is used to simplify certain expressions involving $\frac{e_t}{v_t-2}$ as in (17) below. By (11) and (14) we have

$$v_t - 2 = (r-2) \frac{\tau^t - 1}{\tau - 1} = \frac{\tau^t - 1}{\lambda}. \quad (16)$$

Hence, using (10) and the above relations,

$$\frac{e_t}{v_t - 2} = \frac{\tau^t}{\frac{\tau^t - 1}{\lambda}} = \lambda(1 + c_t) = \lambda + \frac{\tau - 1}{r-2} \frac{1}{\tau^t - 1} = \lambda + \frac{1}{v_t - 2}. \quad (17)$$

Equation (17) is used throughout this section to show that F_t is the sparsest minimal graph that adds e_0 to the graph in t time steps of the K_r -bootstrap process.

Let us recall the following Witness-Set Algorithm introduced in [5]. Given a graph G , we assign a subgraph $F = F(e) \subseteq G$ to each edge $e \in \langle G \rangle_{K_r}$ as follows:

1. If $e \in G$ then set $F(e) = \{e\}$.
2. Choose an order in which to add the edges of $\langle G \rangle_{K_r}$, and at each step identify which r -clique was completed (if more than one is completed then choose one).

3. Add the edges one by one. If e is added by the r -clique K , then set

$$F(e) := \bigcup_{e \neq e' \in K} F(e').$$

A graph F is an r -witness set if there exists a graph G , an edge e , and a realization of the Witness-Set Algorithm (i.e., a sequence of choices as in Step 2) such that $F = F(e)$. The following highly nontrivial extremal result occurs as Lemma 9 in [5], which is stated here without repeating the proof.

Lemma 3.2. Let F be a graph and $r \geq 4$, and suppose that F is an r -witness set. Then,

$$|E(F)| \geq \lambda(|V(F)| - 2) + 1.$$

We say that a graph G is a *minimal graph adding e* if $e \in \langle G \rangle_{K_r}$ but for all proper subgraphs $G' \subsetneq G$ of G we have $e \notin \langle G' \rangle_{K_r}$. It's an immediate observation that every minimal graph adding e to G is an r -witness set. Hence we have the following corollary.

Corollary 3.3. Let $r \geq 4$ and let F be a minimal graph adding e to the graph for some $e \in \langle F \rangle_{K_r}$. Then

$$|E(F)| \geq \lambda(|V(F)| - 2) + 1. \quad (18)$$

We now show that F_t maximises the number of vertices among all minimal graphs that add e_0 to the graph in exactly t time steps of the K_r -bootstrap process.

Lemma 3.4. Let $r \geq 4$, $t \geq 1$ and let F be a minimal graph adding e_0 to the graph at time t in the K_r -bootstrap process. Then $|V(F)| \leq v_t = (r-2)\frac{\tau^t-1}{\tau-1} + 2$ and $|E(F)| \leq e_t = \tau^t$.

Proof. We prove the lemma by induction on t . For $t = 1$ the lemma is trivial as $K_r - e_0$ is the only minimal graph adding e_0 at the first time step. Hence assume that the lemma holds for some $t \geq 1$ and consider a minimal graph F such that e_0 is added at time $t + 1$ in the K_r -bootstrap process started from F .

After one step of the process we obtain a graph F' containing some minimal subgraph F'' that adds e_0 in t additional time steps. By induction we have $|V(F'')| \leq (r-2)\frac{\tau^t-1}{\tau-1} + 2$ and $|E(F'')| \leq \tau^t$. Now, since F was a minimal graph adding e_0 in time $t + 1$, to maximise the number of vertices and edges in F we should in the first step of the process add every edge e of F'' using a copy of $K_r - e$ disjoint from the copies adding other edges in F'' . This shows that $|E(F)| \leq \tau|E(F'')|$ and $|V(F)| \leq |V(F'')| + (r-2)|E(F'')|$. This completes the induction and the lemma follows. \square

A proof closely following that of Lemma 3.4 immediately shows a further extremal result.

Corollary 3.5. For any $t \geq 1$, up to isomorphism, F_t is the only minimal graph on v_t vertices adding e_0 to the graph in exactly t time steps.

As usual, for any graph G and $A, B \subseteq V(G)$, let $E(A, B) = \{\{a, b\} \in E(G) : a \in A, b \in B\}$ and $e(A, B) = |E(A, B)|$. We shall often take $B = V(G)$ and, to simplify the notation, we shall write $e(A, G)$ to denote $e(A, V(G))$. Note that $e(A, G)$ is the number of edges of G with at least one endpoint in A .

In the proofs to come, results on edge-densities of subsets of the graphs $\{F_t : t \geq 1\}$ are proved by induction on t . To make the notation clearer, let us use $E_t(A, B)$ to denote the edges between A and B in the graph F_t and $e_t(A, B) = |E_t(A, B)|$. As usual, $\delta(G) = \min\{\deg_G(v) : v \in V(G)\}$ is used for the minimum degree of G . We shall find the following simple estimate useful in our examination of F_t .

Lemma 3.6. For any $t \geq 2$ and any set $L \subseteq V(F_t)$

$$e_t(L, F_t) \geq \frac{r-1}{2}|L|.$$

Proof. Note that for $t \geq 2$, $\delta(F_t) = r - 1$. Thus

$$(r - 1)|L| \leq \sum_{v \in L} \deg_{F_t}(v) = 2e_t(L, L) + e_t(L, L^c) \leq 2e_t(L, F_t).$$

□

3.1. Overlapping copies of F_t

To prove Statement (i) of Theorem 1.1 we shall show that if p is large enough then with high probability there is a copy of F_t anchored on every pair of vertices in $G_{n,p}$. Towards this aim, we shall show that a measure of the variance of the number of such copies of F_t anchored on a fixed edge e_0 is not too large compared to their expected number. In Section 4, this fact together with Janson's inequality (see Theorem 4.1) is used to deduce the desired result. Hence, we need to prove that it is significantly “harder” (in terms of the ratio of the number of edges to the number of vertices) to find two different such copies of F_t that overlap in at least one vertex (other than $1, 2 \in e_0$) than it is to find two disjoint such copies.

In particular, as the main result in this subsection, it is shown that for any $L \subseteq V(F_t) \setminus \{1, 2\}$,

$$\frac{e_t(L, F_t)}{|L|} \geq \frac{e_t}{v_t - 2} \quad (19)$$

with equality only when $L = V(F_t) \setminus \{1, 2\}$. With this in mind, define ε_t to be such that

$$1 + \varepsilon_t = \left(\frac{v_t - 2}{e_t} \right) \min \left\{ \frac{e_t(L, F_t)}{|L|} : L \subsetneq V(F_t) \setminus \{1, 2\} \right\}. \quad (20)$$

From the definition above, there is no guarantee that ε_t is non-negative. Using induction on t , we shall prove that this is the case by first giving a weak upper bound on ε_t in Lemma 3.7 and then using it to prove a relatively sharp lower bound on ε_t for all $t \geq 1$.

Lemma 3.7. For all $r \geq 4$ and $t \geq 1$ we have $\varepsilon_t \leq \frac{1}{r+1}$.

Proof. First consider the case $t = 1$. For all $L \subsetneq V(F_1) \setminus \{1, 2\}$ the vertices in $F_1 \setminus L$ induce a clique minus the edge joining 1 and 2. Hence, for $1 \leq \ell \leq r - 3$ and $|L| = \ell$, the vertices in $F_1 \setminus L$ induce $\binom{r-\ell}{2} - 1$ edges, which gives

$$\begin{aligned} \frac{e_1(L, F_1)}{|L|} &= \frac{1}{\ell} \left(\binom{r}{2} - 1 - \binom{r-\ell}{2} + 1 \right) \\ &= \frac{1}{2\ell} (r^2 - r - r^2 + 2r\ell - \ell^2 + r - \ell) \\ &= \frac{2r - \ell - 1}{2} \\ &\geq \frac{r+2}{2}, \end{aligned}$$

with equality for $\ell = r - 3$. Hence, by (20)

$$\begin{aligned} \varepsilon_1 &= \left(\frac{r+2}{2} \right) \left(\frac{v_1 - 2}{e_1} \right) - 1 \\ &= \left(\frac{r+2}{2} \right) \left(\frac{r-2}{\binom{r}{2} - 1} \right) - 1 \end{aligned}$$

$$\begin{aligned}
&= (r+2) \left(\frac{r-2}{r^2-r-2} \right) - 1 \\
&= \frac{r+2}{r+1} - 1 = \frac{1}{r+1},
\end{aligned}$$

which proves the lemma for $t = 1$.

Consider now the case when $t \geq 2$. By the construction of the graph F_t , there are vertices in $V(F_t) \setminus \{1, 2\}$ connected to 1 and not to 2. Let v be such a vertex and choose $L = V(F_t) \setminus \{1, 2, v\}$ so that $|L| = v_t - 3$ and $e_t(L, F_t) = e_t - 1$ (recall that $e_t(L, F_t)$ counts all edges in F_t with at least one end in L , thus here it only misses $\{1, v\}$). Hence, for $t \geq 2$ and $r \geq 4$, from the definition in (20) we have

$$\begin{aligned}
\varepsilon_t &\leq \left(\frac{v_t - 2}{e_t} \right) \left(\frac{e_t - 1}{v_t - 3} \right) - 1 \\
&= \frac{e_t v_t - 2e_t - v_t + 2 - e_t v_t + 3e_t}{e_t(v_t - 3)} \\
&= \frac{e_t - v_t + 2}{e_t(v_t - 3)} \\
&< \frac{1}{v_t - 3} \leq \frac{1}{v_2 - 3} \\
&= \frac{1}{(r-2)\binom{r}{2} - 1} \quad (\text{by (11) and (13)}) \\
&\leq \frac{1}{(r-2)6 - 1} \leq \frac{1}{r+1}.
\end{aligned}$$

This completes the proof of Lemma 3.7. \square

The following lemma gives us another result in a similar direction and is used in this section to show that one need only consider certain choices for $L \subseteq V(F_t)$ in order to determine ε_t .

Lemma 3.8. For all $r \geq 4$ and $t \geq 2$ we have

$$\min \left\{ \frac{e_t(L, F_t)}{|L|} : L \subsetneq V(F_t) \setminus \{1, 2\} \right\} < \frac{r+1}{2}.$$

Proof. We prove the lemma by giving an example of a simple set $L \subsetneq V(F_t) \setminus \{1, 2\}$ that satisfies the inequality. Let $v \in V(F_{t-1}) \setminus V(F_{t-2})$. Then, let

$$L = \{v\} \cup \{u \in V(F_t) \setminus V(F_{t-1}) : \{v, u\} \in E(F_t)\}.$$

Since $v \in V(F_{t-1}) \setminus V(F_{t-2})$, we have $\deg_{F_{t-1}}(v) = r - 1$. In F_t , every edge of F_{t-1} is replaced with a copy of K_r minus that edge. Hence for every edge adjacent to v in F_{t-1} there are $r - 2$ neighbours of v in F_t . This implies

$$|L| = 1 + \deg_{F_t}(v) = 1 + (r - 2) \deg_{F_{t-1}}(v) = 1 + (r - 1)(r - 2).$$

To find $e_t(L, F_t)$, i.e., the number of edges adjacent to at least one vertex in L , we notice that all these edges are contained in the $\deg_{F_{t-1}}(v) = r - 1$ copies of K_r minus an edge, that are placed on the edges adjacent to v in F_{t-1} to construct the graph F_t . Therefore we have

$$e_t(L, F_t) = (r - 1) \left(\binom{r}{2} - 1 \right) = \frac{(r - 1)(r + 1)(r - 2)}{2}.$$

This implies that

$$\frac{e_t(L, F_t)}{|L|} = \frac{(r + 1)}{2} \frac{(r - 1)(r - 2)}{1 + (r - 1)(r - 2)} < \frac{(r + 1)}{2}.$$

\square

We observe that, in general, Lemma 3.8 yields a worse upper bound on ε_t than that given by Lemma 3.7, but its form is useful in the proof of Lemma 3.11 to come.

The next theorem is the main tool in the proof of Statement (i) of Theorem 1.1 to come. Here we give a lower bound on ε_t that holds for all $t \geq 1$.

Theorem 3.9. For all $r \geq 4$ and $t \geq 1$,

$$\varepsilon_t \geq \frac{1}{r+1} \left(\frac{2}{r^2-2} \right)^{t-1}. \quad (21)$$

Before proving Theorem 3.9, a few auxiliary lemmas are stated and proved below, along with an outline of the proof. These are then used to establish Theorem 3.9 by induction on t .

Recall that the graph F_{t+1} is constructed by placing an independent copy of $K_r - e$ on every edge e of F_t . Further recall that for each $e \in E(F_t)$ we write $V(e)$ to denote this new set of $r-2$ vertices.

For the induction step from t to $(t+1)$ we will fix a set $L_t \subseteq V(F_t) \setminus \{1, 2\}$ and look for the smallest possible edge densities $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}$ among sets of the form $L_t \cup M$ where $M \subseteq V(F_{t+1}) \setminus V(F_t)$ (see Figure 3).

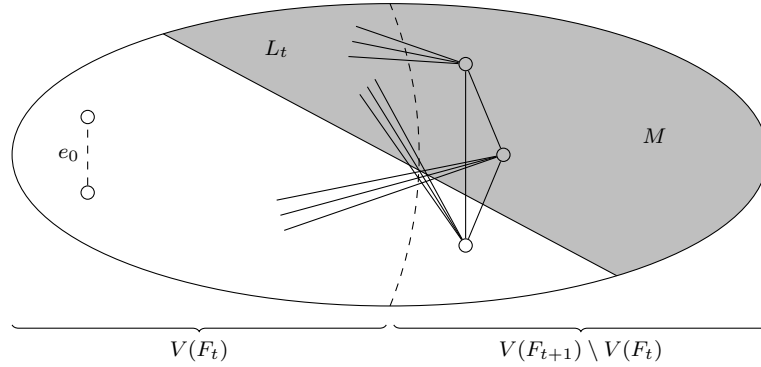


FIG 3. Sets L_t and M in F_{t+1} together with the edges counted in $e_{t+1}(L_t \cup M, F_{t+1})$.

In the following lemma we first deal with the case $L_t = \emptyset$, showing that no set contained entirely in $V(F_{t+1}) \setminus V(F_t)$ can minimise the edge density.

Lemma 3.10. For all $r \geq 4$ and $t \geq 2$ we have

$$\min \left\{ \frac{e_t(L, F_t)}{|L|} : L \subseteq V(F_t) \setminus V(F_{t-1}) \right\} \geq \frac{r+1}{2}.$$

Proof. Let $L \subseteq V(F_t) \setminus V(F_{t-1})$. Hence for every $v \in L$ we have $\deg_{F_t}(v) = r-1$ and at most $r-3$ neighbours of v are also in L . Thus $e_t(L, F_t)$, the number of edges incident to at least one vertex in L , satisfies

$$e_t(L, F_t) \geq |L| \left(\frac{r-3}{2} + 2 \right) = |L| \frac{r+1}{2}$$

and the lemma follows. \square

Thus, by Lemma 3.8, the minimum in equation (20) is not attained for any $L \subseteq V(F_t) \setminus V(F_{t-1})$. We now show that if for some $e \in E(F_t)$ certain conditions are fulfilled, then moving all vertices from $V(e)$ into M does not increase the edge-density minimised in equation (20). The details of this are given in the following lemma.

Lemma 3.11. Let $t \geq 1$, $L_t \subseteq V(F_t) \setminus \{1, 2\}$ with $L_t \neq \emptyset$, and $M \subseteq V(F_{t+1}) \setminus V(F_t)$. For every edge $e = \{x, y\} \in E(F_t)$ such that at least one of x, y is in L_t we have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{e_{t+1}(L_t \cup M \cup V(e), F_{t+1})}{|L_t \cup M \cup V(e)|}.$$

Proof. Let us recall that, by the construction of F_{t+1} , all edges incident to vertices in $V(e)$ are either induced by $V(e)$ or they connect the vertices of $V(e)$ to either x or y .

Set $m = |V(e) \setminus M|$ and note that the conclusion is trivially true for $m = 0$. Thus, assume that $1 \leq m \leq r - 2$.

We consider two different cases. Suppose first that both $x, y \in L_t$. Then $e_{t+1}(L_t \cup M, F_{t+1})$ already counts all edges incident to a vertex in $V(e) \setminus M$, except those that are induced by this set (because the remaining edges are either incident to $x, y \in L_t$ or to some vertex in $V(e) \cap M$). Since $V(e)$ induces a clique in F_{t+1} , we have $e_{t+1}(L_t \cup M \cup V(e), F_{t+1}) - e_{t+1}(L_t \cup M, F_{t+1}) = \binom{m}{2}$. Therefore

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M \cup V(e), F_{t+1})}{|L_t \cup M \cup V(e)|} &= \frac{e_{t+1}(L_t \cup M, F_{t+1}) + \binom{m}{2}}{|L_t \cup M| + m} \\ &= \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m} + \frac{\binom{m}{2}}{m} \frac{m}{|L_t \cup M| + m} \\ &= \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m} + \frac{m-1}{2} \frac{m}{|L_t \cup M| + m} \\ &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m} + \frac{r-3}{2} \frac{m}{|L_t \cup M| + m}. \end{aligned}$$

The above expression is a convex combination of $\frac{r-3}{2}$ and $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}$. By Lemma 3.6, we have $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{r-1}{2} > \frac{r-3}{2}$ and so

$$\frac{e_{t+1}(L_t \cup M \cup V(e), F_{t+1})}{|L_t \cup M \cup V(e)|} \leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}.$$

The case when $x \in L_t$ and $y \notin L_t$ is similar. In this case, $e_{t+1}(L_t \cup M, F_{t+1})$ counts every edge incident to a vertex in $V(e) \setminus M$ that is neither induced by this set nor connects a vertex in this set to y . Since $\{y\} \cup V(e)$ induces a clique in F_{t+1} , we have $e_{t+1}(L_t \cup M \cup V(e), F_{t+1}) - e_{t+1}(L_t \cup M, F_{t+1}) = \binom{m+1}{2}$, and thus

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M \cup V(e), F_{t+1})}{|L_t \cup M \cup V(e)|} &= \frac{e_{t+1}(L_t \cup M, F_{t+1}) + \binom{m+1}{2}}{|L_t \cup M| + m} \\ &= \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m} + \frac{m+1}{2} \frac{m}{|L_t \cup M| + m} \\ &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m} + \frac{r-1}{2} \frac{m}{|L_t \cup M| + m} \\ &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}. \quad (\text{again, by Lemma 3.6}) \end{aligned}$$

The case $x \notin L_t$ and $y \in L_t$ is analogous and this completes the proof. \square

On the other hand, when the edge e does not satisfy the conditions in Lemma 3.11, the following lemma holds.

Lemma 3.12. Let $t \geq 1$, $L_t \subseteq V(F_t) \setminus \{1, 2\}$ with $L_t \neq \emptyset$, $e = \{x, y\} \in E(F_t)$ such that $x, y \notin L_t$, and $M \subseteq V(F_{t+1}) \setminus V(F_t)$. If $|V(e) \cap M| > 0$ then

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} > \min \left\{ \frac{e_{t+1}(L, F_{t+1})}{|L|} : L \subsetneq V(F_{t+1}) \setminus \{1, 2\} \right\}.$$

Proof. Set $m = |V(e) \cap M| > 0$ and recall that we have $m \leq r - 2$. Since $L_t \neq \emptyset$, we also have $L_t \cup M \setminus V(e) \neq \emptyset$. The aim is, by considering the set $L_t \cup M \setminus V(e)$, to show that the set $L_t \cup M$ does not attain the minimum in equation (20).

The edges incident to $V(e) \cap M$ in F_{t+1} can be divided into three groups: those induced by $V(e) \cap M$, those connecting a vertex in $V(e) \cap M$ to a vertex in $V(e) \setminus M$, and those connecting a vertex in $V(e) \cap M$ to either x or y . In total, we have

$$\binom{m}{2} + m(r - 2 - m) + 2m = \binom{m}{2} + m(r - m)$$

edges incident to $V(e) \cap M$ in F_{t+1} . All these edges are clearly counted by $e_{t+1}(L_t \cup M, F_{t+1})$, but not by $e_{t+1}(L_t \cup M \setminus V(e), F_{t+1})$ because $x, y \notin L_t$ and $(L_t \cup M \setminus V(e)) \cap (V(e) \cap M) = \emptyset$. Thus,

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} &= \frac{e_{t+1}(L_t \cup M \setminus V(e), F_{t+1}) + \binom{m}{2} + m(r - m)}{|L_t \cup M \setminus V(e)| + m} \\ &= \frac{e_{t+1}(L_t \cup M \setminus V(e), F_{t+1})}{|L_t \cup M \setminus V(e)|} \cdot \frac{|L_t \cup M \setminus V(e)|}{|L_t \cup M \setminus V(e)| + m} \\ &\quad + \frac{\binom{m}{2} + m(r - m)}{m} \cdot \frac{m}{|L_t \cup M \setminus V(e)| + m} \\ &= \frac{e_{t+1}(L_t \cup M \setminus V(e), F_{t+1})}{|L_t \cup M \setminus V(e)|} \cdot \frac{|L_t \cup M \setminus V(e)|}{|L_t \cup M \setminus V(e)| + m} \\ &\quad + \left(r - \frac{m+1}{2} \right) \cdot \frac{m}{|L_t \cup M \setminus V(e)| + m} \\ &\geq \frac{e_{t+1}(L_t \cup M \setminus V(e), F_{t+1})}{|L_t \cup M \setminus V(e)|} \cdot \frac{|L_t \cup M \setminus V(e)|}{|L_t \cup M \setminus V(e)| + m} \\ &\quad + \frac{r+1}{2} \cdot \frac{m}{|L_t \cup M \setminus V(e)| + m}. \end{aligned}$$

If $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{r+1}{2}$ then the claim holds immediately by Lemma 3.8. Otherwise we have that

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} > \frac{e_{t+1}(L_t \cup M \setminus V(e), F_{t+1})}{|L_t \cup M \setminus V(e)|}.$$

This completes the proof. \square

Recall that by Lemma 3.10, any set $L_t \cup M$ that minimises the ratio $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}$ has $L_t \neq \emptyset$. Furthermore, Lemma 3.12 tells us that any set of the form $L_t \cup M$ with $L_t \neq \emptyset$ that minimises the ratio $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}$ has $|V(e) \cap M| = 0$ for every edge $e \in E_t(L_t^c, L_t^c)$. Let us fix $L_t \subseteq V(F_t) \setminus \{1, 2\}$ and take M to be maximal such that $L_t \cup M$ minimises the edge density.

Assume first that $L_t \subsetneq V(F_t) \setminus \{1, 2\}$ with $L_t \neq \emptyset$. By Lemma 3.11 we see that we then have $|M| = (r-2)e_t(L_t, F_t)$. Also, let e be an arbitrary edge incident to a vertex $x \in L_t$ in F_{t+1} . Again by Lemma 3.11 and by the maximality of M we see that e connects x to a vertex in M : since we have $x \in L_t$, for every edge f adjacent to x in F_t we set $V(f) \subseteq M$, and then e connects x to a vertex in some $V(f)$. Consequently, all edges incident to $L_t \cup M$ in F_{t+1} are incident to M , and since M is a

union of $e_t(L_t, F_t)$ disjoint cliques on $r - 2$ vertices each, such that every vertex in M has exactly two neighbours outside M , we have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} = \frac{e_t(L_t, F_t) \left(\binom{r-2}{2} + 2(r-2) \right)}{|L_t| + (r-2)e_t(L_t, F_t)} = \frac{e_t(L_t, F_t) \left(\binom{r}{2} - 1 \right)}{|L_t| + (r-2)e_t(L_t, F_t)}. \quad (22)$$

Note that when $L_t = V(F_t) \setminus \{1, 2\}$ this choice of M would result in having $L_t \cup M = V(F_{t+1}) \setminus \{1, 2\}$, i.e., the edge density with $L_t = V(F_t) \setminus \{1, 2\}$ is minimised among all choices of M by taking the whole graph. As we want to minimise among all possible proper subsets of $V(F_{t+1})$, the case $L_t = V(F_t) \setminus \{1, 2\}$ requires some further consideration.

We first consider edges $e \in E(F_t)$ that are incident to a vertex in $\{1, 2\}$. For these edges, in the following lemma we show that if, in F_{t+1} , at least one vertex of $V(e)$ is in M , then the edge density does not go up if we move all but one vertex of $V(e)$ to M .

Lemma 3.13. Let $t \geq 1$, $L_t = V(F_t) \setminus \{1, 2\}$ and $M \subsetneq V(F_{t+1}) \setminus V(F_t)$. Let $e = \{x, y\} \in E(F_t)$ be such that $x \in L_t$, $y \in \{1, 2\}$, and let $1 \leq |V(e) \cap M| \leq r - 3$ with $w \in V(e) \setminus M$. We have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{e_{t+1}(L_t \cup M \cup (V(e) \setminus \{w\}), F_{t+1})}{|L_t \cup M \cup V(e) \setminus \{w\}|}.$$

Proof. We prove this lemma analogously to the second part of the proof of Lemma 3.11. Set $m = |V(e) \setminus M| \leq r - 3$.

We have $x \in L_t$ and $y \notin L_t$. In this case, $e_{t+1}(L_t \cup M, F_{t+1})$ counts every edge adjacent to a vertex in $V(e) \setminus (M \cup \{w\})$ except for those that are induced by this set, and those that connect a vertex in this set to either y or w . Since there are $\binom{m-1}{2} + 2(m-1)$ such edges not counted by $e_{t+1}(L_t \cup M, F_{t+1})$, we have

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M \cup (V(e) \setminus \{w\}), F_{t+1})}{|L_t \cup M \cup V(e) \setminus \{w\}|} &= \frac{e_{t+1}(L_t \cup M, F_{t+1}) + \binom{m-1}{2} + 2(m-1)}{|L_t \cup M| + m - 1} \\ &= \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m - 1} \\ &\quad + \frac{m+2}{2} \frac{m-1}{|L_t \cup M| + m - 1}. \end{aligned}$$

Since we have $m \leq r - 3$, we further obtain

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M \cup (V(e) \setminus \{w\}), F_{t+1})}{|L_t \cup M \cup V(e) \setminus \{w\}|} &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m - 1} \\ &\quad + \frac{r-1}{2} \frac{m-1}{|L_t \cup M| + m - 1} \\ &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}. \quad (\text{by Lemma 3.6}) \end{aligned}$$

□

We now turn our attention to edges $e \in E(F_t)$ that are induced by vertices in L_t . Let us show that for them, the edge density does not go up if we move all but one vertex of $V(e)$ to M , even if we initially have $V(e) \cap M = \emptyset$.

Lemma 3.14. Let $t \geq 1$, $L_t = V(F_t) \setminus \{1, 2\}$ with $M \subsetneq V(F_{t+1}) \setminus V(F_t)$. Let $e = \{x, y\} \in E(F_t)$ be such that $x, y \in L_t$, and let $|V(e) \cap M| \leq r - 3$ with $w \in V(e) \setminus M$. We have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{e_{t+1}(L_t \cup M \cup (V(e) \setminus \{w\}), F_{t+1})}{|L_t \cup M \cup V(e) \setminus \{w\}|}.$$

Proof. The proof of this lemma is analogous to the first part of the proof of Lemma 3.11. Set $m = |V(e) \setminus M| \leq r - 2$.

We have $x, y \in L_t$. In this case, $e_{t+1}(L_t \cup M, F_{t+1})$ counts every edge incident to a vertex in $V(e) \setminus (M \cup \{w\})$ except for those that are induced by this set or else connect a vertex in this set to w . Since there are $\binom{m-1}{2} + m - 1$ such edges, we have

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M \cup (V(e) \setminus \{w\}), F_{t+1})}{|L_t \cup M \cup V(e) \setminus \{w\}|} &= \frac{e_{t+1}(L_t \cup M, F_{t+1}) + \binom{m-1}{2} + m - 1}{|L_t \cup M| + m - 1} \\ &= \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m - 1} \\ &\quad + \frac{m}{2} \frac{m - 1}{|L_t \cup M| + m - 1} \\ &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \cdot \frac{|L_t \cup M|}{|L_t \cup M| + m - 1} \\ &\quad + \frac{r - 2}{2} \frac{m - 1}{|L_t \cup M| + m - 1} \\ &\leq \frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}. \quad (\text{by Lemma 3.6}) \end{aligned}$$

□

Consider now the implications of Lemmas 3.13 and 3.14 for the case that $L_t = V(F_t) \setminus \{1, 2\}$ (note that in this case for all $e \in E(F_t)$ we have $e \in E_t(L_t, F_t)$). Let us again take M to be maximal such that $L_t \cup M$ minimises the ratio $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}$.

Since M is maximal, we use Lemma 3.11 to show that there is exactly one edge $e \in E(F_t)$ such that $|V(e) \cap M| < r - 2$. We then use Lemmas 3.13 and 3.14. If e is incident to the vertex 1 (or the vertex 2), then we have either $|V(e) \cap M| = 0$ or $|V(e) \cap M| = r - 3$. If it is not incident to neither 1 nor 2 then we have $|V(e) \cap M| = r - 3$.

In the former case, the vertices in $V(e) \cup \{1\}$ (or $V(e) \cup \{2\}$, respectively) induce a clique on $r - 1$ vertices. The edges of this clique are the only ones not counted by $e_{t+1}(L_t \cup M, F_{t+1})$. Since $V(F_{t+1}) \setminus (L_t \cup M) = \{1, 2\} \cup V(e)$, this implies that we have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{e_{t+1} - \binom{r-1}{2}}{v_{t+1} - r}. \quad (23)$$

In the latter case, if $|V(e) \cap M| = r - 3$, let $\{w\} = V(e) \setminus M$. Since $V(F_{t+1}) \setminus (L_t \cup M) = \{1, 2, w\}$, w can have at most one neighbour, i.e., either vertex 1 or 2, not in $L_t \cup M$. This implies that we have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} \geq \frac{e_{t+1} - 1}{v_{t+1} - 3}. \quad (24)$$

We are now ready to prove Theorem 3.9.

Proof of Theorem 3.9. The proof proceeds by induction on t . Recall the definition of ε_t given in equation (20). Our induction hypothesis is the inequality

$$\varepsilon_t \geq \frac{1}{r+1} \left(\frac{2}{r^2 - 2} \right)^{t-1}$$

(see inequality (21)). We have already seen in the proof of Lemma 3.7 that $\varepsilon_1 = \frac{1}{r+1}$ so the claim holds for $t = 1$. Thus assume the statement is true for some $t \geq 1$. We now proceed with establishing

a recursive lower bound on $e_{t+1}(L_{t+1}, F_{t+1})/|L_{t+1}|$. As before, we consider sets of the form $L_{t+1} = L_t \cup M$, where $L_t \subseteq V(F_t) \setminus \{1, 2\}$ and $M \subseteq V(F_{t+1}) \setminus V(F_t)$, and we write $\ell = |L_t|$. Having established a lower bound on $e_{t+1}(L_{t+1}, F_{t+1})/|L_{t+1}|$ we use (20) to bound ε_t from below. We analyse three cases, depending on the structure of L_t and M , and we show that the first of these cases gives the weakest bound on ε_t from which the theorem follows.

Thus, if there exists a set $L_t \cup M$ that minimises the ratio $\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|}$ for which we have $L_t \neq V(F_t) \setminus \{1, 2\}$ then by (22), we have

$$\frac{e_{t+1}(L_t \cup M, F_{t+1})}{\ell + |M|} \geq \frac{\left(\binom{r}{2} - 1\right) e_t(L_t, F_t)}{\ell + (r-2)e_t(L_t, F_t)} = \frac{\frac{(r-2)(r+1)}{2} e_t(L_t, F_t)}{\ell + (r-2)e_t(L_t, F_t)}. \quad (25)$$

The right-hand side of inequality (25) is a function of the type $\frac{ax}{b+cx}$ with $a, b, c > 0$. Hence it is increasing in x and therefore we can use relation (20) to bound it from below. Thus we have

$$\begin{aligned} \frac{e_{t+1}(L_t \cup M, F_{t+1})}{\ell + |M|} &\geq \frac{\frac{(r-2)(r+1)}{2} \ell (1 + \varepsilon_t) \frac{e_t}{v_t-2}}{\ell + (r-2) \ell (1 + \varepsilon_t) \frac{e_t}{v_t-2}} \\ &\geq \frac{\frac{(r-2)(r+1)}{2} \ell \lambda (1 + c_t) (1 + \varepsilon_t)}{\ell + (r-2) \ell \lambda (1 + c_t) (1 + \varepsilon_t)} \quad (\text{by (17)}) \\ &= \lambda (1 + c_{t+1}) \frac{\frac{(r-2)(r+1)}{2} \frac{(1+c_t)}{(1+c_{t+1})} (1 + \varepsilon_t)}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)}. \end{aligned}$$

Hence, by (20) we see that ε_{t+1} can be bounded from below by

$$\begin{aligned} \varepsilon_{t+1} &\geq \frac{\frac{(r-2)(r+1)}{2} \frac{(1+c_t)}{(1+c_{t+1})} (1 + \varepsilon_t)}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)} - 1 \\ &= \frac{(1 + c_t) (1 + \varepsilon_t) \left(\frac{(r-2)(r+1)}{2} \frac{1}{1+c_{t+1}} - (r-2) \lambda \right) - 1}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)} \\ &= \frac{(1 + c_t) (1 + \varepsilon_t) \left(\frac{\tau}{1+c_{t+1}} - \tau + 1 \right) - 1}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)} \quad (\text{by (13) and (14)}) \\ &= \frac{(1 + c_t) (1 + \varepsilon_t) \left(1 - \frac{1}{\tau^t} \right) - 1}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)} \quad (\text{by (15)}) \\ &= \frac{(1 + c_t) (1 + \varepsilon_t) \frac{1}{1+c_t} - 1}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)} \quad (\text{by (15)}) \\ &= \frac{\varepsilon_t}{1 + (r-2) \lambda (1 + c_t) (1 + \varepsilon_t)}. \quad (26) \end{aligned}$$

Using the bound $\varepsilon_t \leq 1/(r+1)$ in Lemma 3.7, we can bound the expression in (26) from below by

$$\begin{aligned} \varepsilon_{t+1} &\geq \frac{\varepsilon_t}{1 + (r-2) \lambda (1 + c_t) \frac{r+2}{r+1}} \\ &= \frac{\varepsilon_t}{1 + (r-2) \lambda \frac{\tau^t}{\tau^t-1} \frac{r+2}{r+1}} \quad (\text{by (15)}) \\ &= \frac{\varepsilon_t}{1 + (\tau-1) \frac{\tau^t}{\tau^t-1} \frac{r+2}{r+1}} \quad (\text{by (14)}) \\ &\geq \frac{\varepsilon_t}{1 + \tau \frac{r+2}{r+1}} \end{aligned}$$

$$\begin{aligned}
&= \frac{\varepsilon_t}{1 + \frac{(r+2)(r-2)}{2}} && \text{(by (13))} \\
&= \frac{2}{r^2 - 2} \varepsilon_t.
\end{aligned}$$

Thus in this case, by the induction hypothesis, we obtain a lower bound on ε_{t+1} given by

$$\varepsilon_{t+1} \geq \frac{2}{r^2 - 2} \varepsilon_t \geq \frac{1}{r+1} \left(\frac{2}{r^2 - 2} \right)^t.$$

If on the other hand there exists a set $L_t \cup M$ with $L_t = V(F_t) \setminus \{1, 2\}$ that gives the minimal ratio in equation (20) then we shall consider two cases as outlined in the discussion before this proof. If $M = V(F_{t+1}) \setminus (V(F_t) \cup V(e))$ for some $e \in E(F_t)$ then by (23) we have

$$\begin{aligned}
\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} &\geq \frac{e_{t+1} - \binom{r-1}{2}}{v_{t+1} - r} && \text{(by (23))} \\
&= \frac{\tau^{t+1} - \binom{r-1}{2}}{(r-2) \frac{(\tau^{t+1}-1)}{\tau-1} - (r-2)} && \text{(by (10) and (11))} \\
&= \frac{e_{t+1}}{v_{t+1} - 2} \left(\frac{1 - \frac{r^2-3r+2}{2\tau^{t+1}}}{1 - \frac{(r-2)(\tau-1)}{(r-2)(\tau^{t+1}-1)}} \right).
\end{aligned}$$

As before, by (20) we see that in this case ε_{t+1} is bounded from below by

$$\begin{aligned}
\varepsilon_{t+1} &\geq \frac{1 - \frac{r^2-3r+2}{2\tau^{t+1}}}{1 - \frac{\tau-1}{\tau^{t+1}-1}} - 1 \\
&= \frac{\frac{\tau-1}{\tau^{t+1}-1} - \frac{r^2-3r+2}{2\tau^{t+1}}}{1 - \frac{\tau-1}{\tau^{t+1}-1}} \\
&= \frac{\tau^{t+2} - \tau^{t+1} - (\tau^{t+1} - 1)(\tau - r + 2)}{\tau^{t+1}(\tau^{t+1} - 1)} \cdot \frac{1}{1 - \frac{\tau-1}{\tau^{t+1}-1}} && \text{(by (13))} \\
&> \frac{(r-3)\tau^{t+1} + (r-2)(\frac{\tau+1}{2} - 1)}{\tau^{t+1}(\tau^{t+1} - 1)} && \text{(by (13))} \\
&> \frac{(r-3)(\tau^{t+1} + 1)}{\tau^{t+1}(\tau^{t+1} - 1)} \\
&> \frac{r-3}{\tau^{t+1}}.
\end{aligned}$$

The last inequality is clearly weaker than it could be and is given in this form to make further calculations simpler.

If $L_t = V(F_t) \setminus \{1, 2\}$ and $M = V(F_{t+1}) \setminus (V(F_t) \cup \{w\})$ for some $w \in V(F_{t+1}) \setminus V(F_t)$ then by (24) we have

$$\begin{aligned}
\frac{e_{t+1}(L_t \cup M, F_{t+1})}{|L_t \cup M|} &\geq \frac{e_{t+1} - 1}{v_{t+1} - 3} && \text{(by (24))} \\
&= \frac{\tau^{t+1} - 1}{(r-2) \frac{(\tau^{t+1}-1)}{\tau-1} - 1} && \text{(by (10) and (11))} \\
&= \frac{e_{t+1}}{v_{t+1} - 2} \left(\frac{1 - \frac{1}{\tau^{t+1}}}{1 - \frac{\tau-1}{(r-2)(\tau^{t+1}-1)}} \right). && \text{(by (16))}
\end{aligned}$$

Thus, we see that by (20) in this case ε_{t+1} is at least

$$\begin{aligned}
\varepsilon_{t+1} &\geq \frac{1 - \frac{1}{\tau^{t+1}}}{1 - \frac{\tau-1}{(r-2)(\tau^{t+1}-1)}} - 1 \\
&= \frac{\frac{\tau-1}{(r-2)(\tau^{t+1}-1)} - \frac{1}{\tau^{t+1}}}{1 - \frac{\tau-1}{(r-2)(\tau^{t+1}-1)}} \\
&= \frac{\tau^{t+2} - (r-1)\tau^{t+1} + (r-2)}{\tau^{t+1}(r-2)(\tau^{t+1}-1)} \cdot \frac{1}{1 - \frac{(\tau-1)}{(r-2)(\tau^{t+1}-1)}} \\
&> \frac{\tau^{t+2} - (r-1)\tau^{t+1}}{\tau^{t+1}(r-2)(\tau^{t+1}-1)} \\
&> \frac{\tau - (r-1)}{(r-2)\tau^{t+1}} \\
&= \frac{1}{\tau^{t+1}} \left(\frac{r+1}{2} - \frac{r-1}{r-2} \right) \quad (\text{by (13)}) \\
&= \frac{1}{\tau^{t+1}} \left(\frac{r+1}{2} - \frac{1}{r-2} - 1 \right) \\
&= \frac{\lambda-1}{\tau^{t+1}} \quad (\text{by (12)}).
\end{aligned}$$

It remains to show that for all $t \geq 0$ we have

$$\frac{1}{r+1} \left(\frac{2}{r^2-2} \right)^t \leq \min \left\{ \frac{r-3}{\tau^{t+1}}, \frac{\lambda-1}{\tau^{t+1}} \right\}. \quad (27)$$

To see that the inequality in (27) holds, note that since $r \geq 4$,

$$\frac{1}{r+1} = \frac{r-2}{(r+1)(r-2)} \leq \frac{\frac{r+1}{2} - \frac{1}{r-2} - 1}{\frac{(r+1)(r-2)}{2}} = \frac{\lambda-1}{\tau},$$

and

$$\frac{1}{r+1} \leq \frac{1}{r+1} \frac{2(r-3)}{r-2} = \frac{r-3}{\tau}.$$

Moreover,

$$\frac{2}{r^2-2} < \frac{2}{r^2-r-2} = \frac{1}{\tau}.$$

Thus, for any $t \geq 0$, the terms in equation (27) satisfy $\frac{1}{r+1} \left(\frac{2}{r^2-2} \right)^t \leq \frac{\lambda-1}{\tau} \cdot \frac{1}{\tau^t}$ and $\frac{1}{r+1} \left(\frac{2}{r^2-2} \right)^t \leq \frac{r-3}{\tau} \cdot \frac{1}{\tau^t}$. This completes the induction on t and the proof of Theorem 3.9. \square

Let us conclude this section by commenting on the sharpness of the bound in (21). We know that $\varepsilon_1 = 1/(r+1)$ is obtained by taking $L = L_1$ of size $r-3$, i.e., by leaving just one vertex in $V(F_1) \setminus \{1, 2\}$ outside L_1 . If we then continue by, for $2 \leq i \leq t$, taking L_i to be the union of L_{i-1} and all vertices in $V(F_i) \setminus V(F_{i-1})$ that are adjacent to at least one vertex in L_{i-1} , then the resulting set L_t has $1 + 2(r-2)\frac{\tau^{t-1}-1}{\tau-1}$ vertices and $\tau^{t-1}(\tau-2)$ edges adjacent to it. This shows that for some $C_r > 0$, one can obtain a bound $\varepsilon_t \leq C_r/\tau^t$. Since $\tau = (r^2 - r - 2)/2$, this implies that our lower bound on ε_t is relatively sharp.

We have thus shown that the edge density of all proper subsets of $V(F_t)$ is strictly bounded from below by $e_t/(v_t - 2)$. As we will see in Section 4, we are now equipped with the necessary means to prove Statement (i) of Theorem 1.1.

4. Upper bound on the critical probability

In this section we prove Statement (i) of Theorem 1.1. We shall use the following form of Janson's inequality.

Theorem 4.1. Let R be a set and let $S \subseteq R$ be a random subset of R , where each $r \in R$ is in S independently with probability p . Let $\{B_1, \dots, B_m\}$ be a collection of finite subsets of R and let C_i be the event that $B_i \subseteq S$. Let $Z = \sum_{i=1}^m \mathbb{1}_{\{C_i\}}$ and let $\mu = \sum_{i=1}^m \mathbb{P}_p(C_i) = \mathbb{E}[Z]$. For $1 \leq i, j \leq m$, $i \neq j$, let $i \sim j$ if $B_i \cap B_j \neq \emptyset$, i.e., if the events C_i and C_j are dependent, and let $\Delta = \sum_{i \sim j} \mathbb{P}(C_i \cap C_j)$. Then

$$\mathbb{P}_p(Z = 0) \leq e^{-\mu + \Delta/2}. \quad (28)$$

□

Let $e_0 = \{1, 2\}$. In this section we show that if $p(n) \geq n^{-\frac{v_t-2}{e_t}} \log n$ then we have $\mathbb{P}_{p_n}(e_0 \notin E(G_t)) \leq n^{-3}$. By the union bound this implies $\mathbb{P}_{p_n}(T \leq t) \rightarrow 1$.

Proof of Statement (i) of Theorem 1.1. Fix n , sufficiently large, and $t = t(n) \leq \frac{\log \log n}{3 \log \tau}$.

As always, fix two vertices 1 and 2 and let $e_0 = \{1, 2\}$. Given any of the $\binom{n-2}{v_t-2}$ subsets $X_i \subseteq \{3, 4, \dots, n\}$ of size $(v_t - 2)$, let $F_t(X_i)$ be an arbitrary fixed copy of the graph F_t on $X_i \cup \{1, 2\}$ that adds the edge e_0 to the graph G in t time steps. We shall apply Theorem 4.1 to bound the probability of e_0 not being added to the graph by time t from above.

Let $p(n) = n^{-(v_t-2)/e_t} \log n$, as in Statement (i) and let $G = G_{n,p(n)}$. In Theorem 4.1 we shall take $R = [n]^{(2)}$, $S = E(G)$ and for $i = 1, \dots, \binom{n-2}{v_t-2}$ let $B_i = E(F_t(X_i))$. We define C_i , as well as Z , as in Theorem 4.1. We clearly have $\mu = \sum_{i=1}^m \mathbb{P}_p(C_i) = \binom{n-2}{v_t-2} p^{e_t}$.

An upper bound on Δ in Theorem 4.1 can be obtained by considering the following form,

$$\Delta = \sum_{i \sim j} \mathbb{P}(C_i \cap C_j) = \sum_{1 \leq |X_i \setminus X_j| \leq v_t-3} \mathbb{P}(B_i \subseteq S \text{ and } B_j \subseteq S).$$

There are $\binom{n-2}{v_t-2}$ ways to choose the set X_i and, having fixed X_i , for $1 \leq k \leq v_t - 3$ there are $\binom{v_t-2}{v_t-2-k} \binom{n-v_t}{k}$ ways to choose the set X_j such that $|X_j \setminus X_i| = k$. We also have $|B_i| = e_t$ and, by (20), $|B_j \setminus B_i| \geq k(1 + \varepsilon_t) \frac{e_t}{v_t-2}$. Hence

$$\begin{aligned} \Delta &\leq \binom{n-2}{v_t-2} \sum_{k=1}^{v_t-3} \binom{v_t-2}{v_t-2-k} \binom{n-v_t}{k} p^{e_t + k(1+\varepsilon_t) \frac{e_t}{v_t-2}} \\ &= \binom{n-2}{v_t-2} p^{e_t} \sum_{k=1}^{v_t-3} \binom{v_t-2}{k} \binom{n-v_t}{k} p^{k(1+\varepsilon_t) \frac{e_t}{v_t-2}} \\ &= \mu \sum_{k=1}^{v_t-3} \binom{v_t-2}{k} \binom{n-v_t}{k} p^{k(1+\varepsilon_t) \frac{e_t}{v_t-2}} \\ &< \mu \sum_{k=1}^{v_t-3} ((v_t-2)n)^k p^{k(1+\varepsilon_t) \frac{e_t}{v_t-2}}. \end{aligned} \quad (29)$$

Note that, by the definition of ε_t in (20), by Lemma 3.8 and (12),

$$\frac{e_t}{v_t-2} (1 + \varepsilon_t) \leq \frac{r+1}{2} = \lambda + \frac{1}{r-2} < 2\lambda. \quad (30)$$

By (14), for all $r \geq 4$ we have $\tau - 1 > r - 2$ hence, by (16), $v_t - 2 < \tau^t$. For $t \leq \frac{\log \log n}{3 \log \tau} = \frac{\log_r(\log n)}{3}$, by (14) we have

$$v_t - 2 < \tau^t \leq (\log n)^{1/3}. \quad (31)$$

Now, using the fact that $p(n) = n^{-(v_t-2)/e_t} \log n$, we have

$$\begin{aligned}
\sum_{k=1}^{v_t-3} ((v_t-2)n)^k p^{k \frac{e_t}{v_t-2}(1+\varepsilon_t)} &= \sum_{k=1}^{v_t-3} ((v_t-2)n)^k n^{-k(1+\varepsilon_t)} (\log n)^{k \frac{e_t}{v_t-2}(1+\varepsilon_t)} \\
&\leq \sum_{k=1}^{v_t-3} ((v_t-2)n)^k (\log n)^{2\lambda k} n^{-k(1+\varepsilon_t)} && \text{(by (30))} \\
&\leq \sum_{k=1}^{v_t-3} (\log n)^{k/3} (\log n)^{2\lambda k} n^{-k(1+\varepsilon_t)} && \text{(by (31))} \\
&= \sum_{k=1}^{v_t-3} (\log n)^{k(2\lambda+1/3)} n^{-k\varepsilon_t}. && (32)
\end{aligned}$$

By inequality (21) in Theorem 3.9, we have $\varepsilon_t \geq \frac{1}{r+1}(2/(r^2-2))^{t-1}$. This implies that for $t \leq \frac{\log \log n}{3 \log \tau}$ we have $\varepsilon_t > (\log n)^{-1/2}$. Indeed,

$$\varepsilon_t \geq \frac{1}{r+1} \left(\frac{2}{r^2-2} \right)^{t-1} > \frac{1}{r+1} \left(\frac{1}{\tau} \frac{2\tau}{r^2-2} \right)^{\frac{\log \log n}{3 \log \tau}} = \frac{1}{r+1} (\log n)^{-1/3} \left(\frac{r^2-r-2}{r^2-2} \right)^{\frac{\log \log n}{3 \log \tau}}.$$

Now, since the sequence $a_r = \frac{r^2-r-2}{r^2-2}$ is increasing in r and for all $r \geq 4$ we have $\tau \geq 5$,

$$\left(\frac{r^2-r-2}{r^2-2} \right)^{\frac{1}{3 \log \tau}} \geq \left(\frac{5}{7} \right)^{\frac{1}{3 \log 5}} > 0.93 > e^{-1/10},$$

and the bound $\varepsilon_t \geq (\log n)^{-1/2}$ follows for n large enough. Consequently, for n large enough we have that $(\log n)^{2\lambda+1/3} n^{-\varepsilon_t} \leq (\log n)^{2\lambda+1/3} \exp(-(\log n)^{1/2}) < 1/2$. Hence continuing the string of inequalities from (29) and (32),

$$\Delta \leq \mu \sum_{k=1}^{v_t-3} \left((\log n)^{2\lambda+1/3} n^{-\varepsilon_t} \right)^k \leq \mu \sum_{k=1}^{\infty} \left(\frac{1}{2} \right)^k = \mu. \quad (33)$$

By (31), we see that the ratio $(\frac{n-2}{n})^{v_t}$ tends to 1 as $n \rightarrow \infty$. Thus, for $p \geq n^{-\frac{v_t-2}{e_t}} \log n$ we have

$$\mu = \binom{n-2}{v_t-2} p^{e_t} \geq \left(\frac{n}{v_t-2} \right)^{v_t-2} n^{-\frac{v_t-2}{e_t} e_t} (\log n)^{e_t} = \frac{(\log n)^{e_t}}{(v_t-2)^{v_t-2}} \geq \left(\frac{\log n}{\tau^t} \right)^{\tau^t}, \quad (34)$$

where the second inequality follows from (10) and the bound on v_t in (31). Hence using Theorem 4.1 we obtain

$$\begin{aligned}
\mathbb{P}(Z=0) &\leq \exp(-\mu + \Delta/2) && \text{(by (28))} \\
&\leq \exp(-\mu/2) && \text{(by (33))} \\
&\leq \exp\left(-\frac{1}{2} \left(\frac{\log n}{\tau^t} \right)^{\tau^t}\right). && \text{(by (34))}
\end{aligned}$$

Note that the function $x \mapsto \left(\frac{\log n}{x} \right)^x$ is increasing for $x \in (0, \frac{\log n}{e}]$. When $t=1$, for n sufficiently large,

$$\left(\frac{\log n}{\tau} \right)^{\tau} = \log n \frac{(\log n)^{\tau-1}}{\tau^{\tau}} \geq 6 \log n.$$

For $t \leq \frac{\log \log n}{3 \log \tau}$ we have $\tau \leq \tau^t \leq (\log n)^{1/3} < \frac{\log n}{e}$ and so,

$$\left(\frac{\log n}{\tau^t}\right)^{\tau^t} \geq \left(\frac{\log n}{\tau}\right)^{\tau} \geq 6 \log n$$

when n is sufficiently large. Thus,

$$\mathbb{P}(Z = 0) \leq \exp\left(-\frac{1}{2} \left(\frac{\log n}{\tau^t}\right)^{\tau^t}\right) \leq \exp(-3 \log n) = n^{-3},$$

and applying the union bound yields

$$\mathbb{P}_{p_n}(T \leq t) \geq 1 - \frac{1}{n}.$$

This completes the proof of Statement (i) of Theorem 1.1. \square

5. Lower bound on the critical probability

In this section we prove Statement (ii) of Theorem 1.1. More precisely, we show that if $p(n) = o\left(n^{-\frac{v_t-2}{e_t}}\right)$ then even a single fixed pair $e_0 = \{1, 2\}$ is not added to the graph by time t with high probability.

Proof of Statement (ii) of Theorem 1.1. Recall that $E(G_t)$ denotes the edges of the graph after t time steps. We have

$$\mathbb{P}_{p_n}(T \leq t) \leq \mathbb{P}_{p_n}(e_0 \in E(G_t)) \leq \sum_{i=0}^t \mathbb{P}_{p_n}(e_0 \text{ is added at time } i).$$

The probability that e_0 is added to the graph at time 0 is clearly $p_n = o(1)$. For $1 \leq i \leq t-1$ we can be very generous with our estimates. The number of vertices of any minimal graph that adds e_0 to the graph at time $i \geq 1$ is at least r (including 1 and 2) and, by Lemma 3.4, at most v_i . The number of different graphs on a set of j vertices is $2^{\binom{j}{2}} < 2^{j^2}$. Finally, by Corollary 3.3 any minimal graph on j vertices that adds e_0 to the graph contains at least $\lambda(j-2) + 1$ edges. Since we take $p(n) = o\left(n^{-\frac{v_t-2}{e_t}}\right)$, we see that by the union bound the probability that e_0 is added to the graph at some time $1 \leq i \leq t-1$ is at most

$$\begin{aligned} \sum_{i=1}^{t-1} \sum_{j=r}^{v_i} \binom{n}{j-2} 2^{\binom{j}{2}} p^{\lambda(j-2)+1} &\leq \sum_{i=1}^{t-1} \sum_{j=r}^{v_i} n^{j-2} 2^{j^2} n^{-(\lambda(j-2)+1) \frac{v_t-2}{e_t}} \\ &= \sum_{i=1}^{t-1} \sum_{j=r-2}^{v_i-2} n^j 2^{(j+2)^2} n^{-(\lambda(j+1) \frac{v_t-2}{e_t})}. \end{aligned} \quad (35)$$

For $i = t$ we further divide into two cases. We deal with the graphs on at most $v_t - 1$ vertices in the same way as we did for $i < t$. By union bound, the probability that any one of them appears in the graph is at most

$$\sum_{j=r-2}^{v_t-3} n^j 2^{(j+2)^2} n^{-(\lambda(j+1) \frac{v_t-2}{e_t})}. \quad (36)$$

The last case we need to consider are graphs on v_t vertices that add e_0 to the graph at time t . By Corollary 3.5 there are at most n^{v_t-2} such graphs and the probability that we obtain at least one of them, for some $\omega(n)$ that tends to infinity as $n \rightarrow \infty$, is at most

$$n^{v_t-2} n^{-e_t \frac{v_t-2}{e_t}} / \omega(n)^{e_t} = \omega(n)^{-e_t} = o(1).$$

Consequently we have

$$\mathbb{P}_{p_n}(T \leq t) \leq \sum_{i=1}^{t-1} \sum_{j=r-2}^{v_i-2} n^j 2^{(j+2)^2} n^{-(\lambda j+1) \frac{v_i-2}{e_i}} + \sum_{j=r-2}^{v_t-3} n^j 2^{(j+2)^2} n^{-(\lambda j+1) \frac{v_t-2}{e_t}} + o(1). \quad (37)$$

For all $t \geq 1$, by (16) we have

$$v_{t-1} - 2 = (r-2) \frac{\tau^{t-1} - 1}{\tau - 1} < \frac{r-2}{\tau} \frac{\tau^t - 1}{\tau - 1} = \frac{v_t - 2}{\tau}.$$

It follows that for $i \leq t-1$ we can bound the powers of n in (35) by

$$\begin{aligned} j - (\lambda j + 1) \frac{(v_t - 2)}{e_t} &= j - j \frac{\lambda(v_t - 2)}{e_t} - \frac{v_t - 2}{e_t} \\ &= j \left(1 - \frac{1}{1 + c_t} \right) - \frac{1}{\lambda(1 + c_t)} \quad (\text{by (17)}) \\ &= \left(jc_t - \frac{1}{\lambda} \right) \frac{1}{1 + c_t} \\ &\leq \left((v_{t-1} - 2)c_t - \frac{1}{\lambda} \right) \frac{1}{1 + c_t} \\ &= \left(\frac{r-2}{\tau-1} (\tau^{t-1} - 1) \frac{1}{\tau^t - 1} - \frac{r-2}{\tau-1} \right) \frac{1}{1 + \frac{1}{\tau^t - 1}} \quad (\text{by (16), (15) and (14)}) \\ &= \frac{r-2}{\tau-1} \frac{\tau^{t-1} - 1}{\tau^t - 1} \frac{\tau^t - 1}{\tau^t} \\ &= -\frac{r-2}{\tau}. \end{aligned}$$

Analogously, for $i = t$ and $j \leq v_t - 3$, we can bound the powers of n in (36) by

$$\begin{aligned} j - (\lambda j + 1) \frac{(v_t - 2)}{e_t} &= \left(jc_t - \frac{1}{\lambda} \right) \frac{1}{1 + c_t} \\ &\leq \left((v_t - 3)c_t - \frac{1}{\lambda} \right) \frac{1}{1 + c_t} \\ &= \left(\frac{\tau^t - 1}{\lambda} \frac{1}{\tau^t - 1} - c_t - \frac{1}{\lambda} \right) \frac{1}{1 + c_t} \quad (\text{by (16) and (15)}) \\ &= \frac{-c_t}{1 + c_t} = \frac{-1/(\tau^t - 1)}{1 + 1/(\tau^t - 1)} \\ &= \frac{-1}{\tau^t}. \end{aligned}$$

We can use these estimates and the fact that, by (16), for any $i \geq 1$ we have $v_i - 2 = \frac{\tau^i - 1}{\lambda} < \tau^i / \lambda$, to bound $\mathbb{P}_{p_n}(T \leq t)$ from above. Indeed,

$$\mathbb{P}_{p_n}(T \leq t) \leq \sum_{i=1}^{t-1} \sum_{j=r-2}^{v_i-2} 2^{(j+2)^2} n^{-(r-2)/\tau} + \sum_{j=r-2}^{v_t-3} 2^{(j+2)^2} n^{-1/\tau^t} + o(1)$$

$$\begin{aligned}
&\leq (t-1)(v_{t-1}-r+1)2^{(v_{t-1})^2}n^{-(r-2)/\tau} + (v_t-r)2^{(v_t)^2}n^{-1/\tau^t} + o(1) \\
&< t\frac{\tau^{t-1}}{\lambda}2^{\tau^{2(t-1)}/\lambda^2}n^{-(r-2)/\tau} + \frac{\tau^t}{\lambda}2^{\tau^{2t}/\lambda^2}n^{-1/\tau^t} + o(1).
\end{aligned} \tag{38}$$

There is some constant $C'_r > 0$ such that for all $t \geq C'_r$ we have

$$2^{\tau^{2(t-1)}/\lambda^2} \geq t\frac{\tau^{t-1}}{\lambda} \quad \text{and} \quad 2^{\tau^{2t}/\lambda^2} \geq \frac{\tau^t}{\lambda}.$$

For $t < C'_r$ all three terms in (38) tend to 0 as $n \rightarrow \infty$ and we clearly have $\mathbb{P}_{p_n}(T \leq t) = o(1)$. For $t \geq C'_r$ we continue (38) to obtain

$$\begin{aligned}
\mathbb{P}_{p_n}(T \leq t) &\leq 2^{2\tau^{2(t-1)}/\lambda^2}n^{-(r-2)/\tau} + 2^{\tau^{2t}/\lambda^2}n^{-1/\tau^t} + o(1) \\
&\leq \exp\left(\frac{2\tau^{2t}}{\lambda^2}\log 2 - \frac{(r-2)}{\tau}\log n\right) + \exp\left(\tau^{2t}\log 2 - \frac{1}{\tau^t}\log n\right) + o(1).
\end{aligned}$$

Thus, for $C'_r \leq t \leq \frac{\log \log n}{3 \log \tau}$, with all logarithms having base e ,

$$\begin{aligned}
\mathbb{P}_{p_n}(T \leq t) &\leq \exp\left(\frac{2(\log n)^{2/3}}{\lambda^2}\log 2 - \frac{(r-2)}{\tau}\log n\right) \\
&\quad + \exp\left((\log n)^{2/3}\log 2 - (\log n)^{2/3}\right) + o(1) \\
&= o(1).
\end{aligned}$$

This completes the proof of Theorem 1.1. □

6. Open problems

In this paper we determine the critical probability for percolation by time t in K_r -bootstrap percolation up to a logarithmic factor. The first obvious problem to consider is the following.

Problem 6.1. Close the gap between Statement (i) and Statement (ii) in Theorem 1.1.

We do not expect neither of our bounds to be sharp. However, we believe that for the range of t discussed in this paper we have $p_c(n, r, t)/n^{-(v_t-2)/e_t} \rightarrow \infty$.

The second open problem we pose here is of extremal nature. Lemma 3.4 tells us that minimal graphs adding e_0 to the graph at time t have at most $(r-2)\frac{\tau^t-1}{\tau-1} + 2$ vertices and τ^t edges.

Problem 6.2. How small, both in terms of the size of the vertex set and the edge set, can minimal graphs adding e_0 at time t be?

Acknowledgement

The authors wish to thank the referees for their very detailed feedback and, in particular, for their many suggestions that helped to improve the presentation of a number of the technical lemmas.

References

- [1] N. Alon, *An extremal problem for sets with applications to graph theory*, J. Combin. Theory, Ser. A **40** (1985), 82–89.
- [2] P. Balister, B. Bollobás, and P. Smith, *The time of bootstrap percolation in two dimensions*, To appear in *Probability Theory and Related Fields*, <http://arxiv.org/abs/1305.5444>.

- [3] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris, *The sharp threshold for bootstrap percolation in all dimensions*, Transactions of the American Mathematical Society **364** (2012), 2667–2701.
- [4] J. Balogh, B. Bollobás, and R. Morris, *Bootstrap percolation in three dimensions*, Annals of Probability **37** (2009), 1329–1380.
- [5] ———, *Graph bootstrap percolation*, Random Structures and Algorithms **41** (2012), 413–440.
- [6] J. Balogh and B. G. Pittel, *Bootstrap percolation on the random regular graph*, Random Structures and Algorithms **30** (2007), 257–286.
- [7] B. Bollobás, *Weakly k -saturated graphs*, In Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967), H. Sachs, H. J. Voss, and H. Walther (Editors), Teubner, Leipzig, Barth, 1968, pp. 25–31.
- [8] ———, *The diameter of random graphs*, Transactions of the American Mathematical Society **267** (1981), 41–52.
- [9] ———, *Graph bootstrap percolation*, (2011), preprint.
- [10] B. Bollobás, K. Gunderson, C. Holmgren, S. Janson, and M. Przykucki, *Bootstrap percolation on Galton–Watson trees*, Electronic Journal of Probability **19** (2014), 1–27.
- [11] B. Bollobás, C. Holmgren, P. Smith, and A.J. Uzzell, *The time of bootstrap percolation with dense initial sets*, Annals of Probability **4** (2014), 1337–1373.
- [12] B. Bollobás, P. Smith, and A.J. Uzzell, *The time of bootstrap percolation with dense initial sets for all thresholds*, To appear in *Random Structures and Algorithms*, <http://arxiv.org/abs/1209.4339>.
- [13] J. Chalupa, P.L. Leath, and G.R. Reich, *Bootstrap percolation on a Bethe lattice*, Journal of Physics C **12** (1979), L31–L35.
- [14] A. van Enter, *Proof of Straley’s Argument for Bootstrap Percolation*, Journal of Statistical Physics **48** (1987), 943–945.
- [15] P. Frankl, *An extremal problem for two families of sets*, Eur. J. Math. **3** (1982), 125–127.
- [16] A. E. Holroyd, *Sharp metastability threshold for two-dimensional bootstrap percolation*, Probability Theory and Related Fields **125** (2003), 195–224.
- [17] S. Janson, T. Łuczak, T. Turova, and T. Vallier, *Bootstrap percolation on the random graph $G_{n,p}$* , Annals of Applied Probability **22** (2012), 1989–2047.
- [18] G. Kalai, *Weakly saturated graphs are rigid*, Convexity and graph theory, M. Rosenfeld and J. Zaks (Editors), Ann. Discrete Math., 1984, pp. 189–190.
- [19] J. von Neumann, *Theory of self-reproducing automata*, University of Illinois Press, 1966.
- [20] R.H. Schonmann, *On the behaviour of some cellular automata related to bootstrap percolation*, Annals of Probability **20** (1992), 174–193.
- [21] S. Ulam, *Random processes and transformations*, Proceedings of the International Congress of Mathematicians, Vol. 2, Cambridge, Massachusetts, August 30–September 6, 1950, 1952, pp. 264–275.